About the intrinsic resistance variability in HfO_2-based RRAM devices

C. Cagli[1], G. Piccolboni[1], D. Garbin[1], A. Grossi[1,2], G. Molas[1], E. Vianello[1], C. Zambelli[2], G. Reimbold[1]
[1]CEA-LETI, Minatec Campus, Grenoble, France
Email: carlo.cagli@cea.fr
[2]ENDIF, Univ. of Ferrara, Ferrara, Italy
Email: alessandro.grossi@unife.it

***Abstract*—Resistive memories (RRAM) are attracting a wide interest as candidates for the next generation memory technology, in particular for embedded and, more generally, for low power applications like IoT. However, their variability still remains a concern as the latter has been proven to be an intrinsic feature linked to the filamentary switching mechanism. In this paper we will review some recent key results by our and other groups that show the fundamental variability limits of the HfO_2-based technology and some possible alternatives.**

I. Introduction

RRAM displays several key features that made this technology a highly ranked candidate for a broad panel of applications. In particular, RRAM characterizes itself for the low working voltage, the high speed and high endurance, and for a good data retention [1-3]. The applications that could best benefit from RRAM features are the embedded technologies (eNVM) and those solutions that require intrinsically a low voltage and low power, such as the IoT (Internet of Things). Moreover the RRAM is still considered a candidate for the so-called Storage Class Memory (SCM) thanks to its performances fitting between the DRAM and the mass storage (Flash and magnetic disks) ones. HfO_2 has been widely investigated as RRAM material and represents the leader candidate. What prevents RRAM from being a more mature technology is certainly the high resistance variability that is generally observed in particularly in the High Resistance State (HRS). In this paper we will show some recent results published by our and other groups that explain the nature of this variability and its limit. The impact on the RRAM is showed and quantified. Eventually some interesting findings and considerations on alternative materials are shown.

This paper is organized in the following way: first, we explain how the cycle-to-cycle (C2C) variability lowers when a restricted amount of cycles are considered, due to a short range correlation. Next, we show the intrinsic relationship between the median HRS value and its spread in terms of standard deviation. This will be first illustrated on a limited number of cells and explained by a simple model, and then the argument will be spread to a larger cell array. At this point the intrinsic variability in HfO_2-based RRAM will be clear. Eventually some results recently published by *A. Bricalli et al.* [4] will be shown, reporting an improvement of the variability limit with a different material, namely SiO_x.

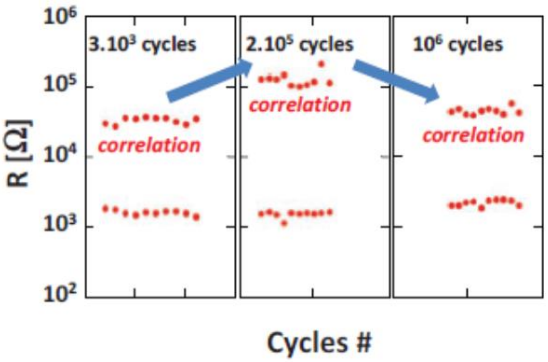


Figure 1 The figure shows the typical short range correlation among cycles. From [5]

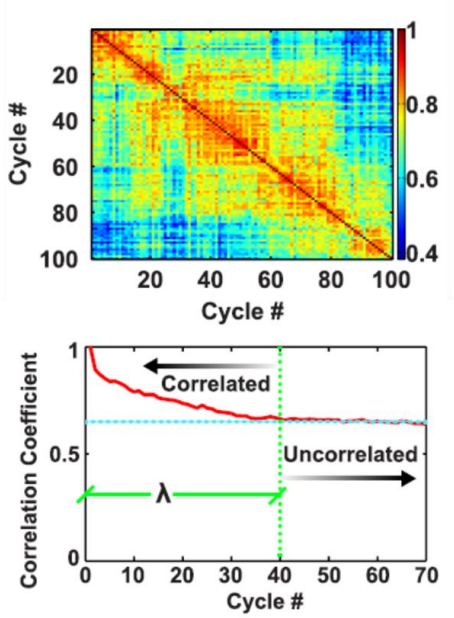


Figure 2 *Up:* correlation of resistances during cycling. *Down:* Horizontal cut showing the values of correlation coefficient. From [5]

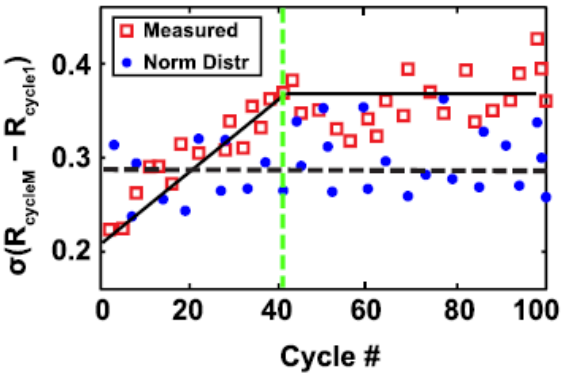


Figure 4 As an effect of internal correlation, σ increases with cycle range (*red*). As a contrast no increase is observed for a random data set (*blue*). From [5]

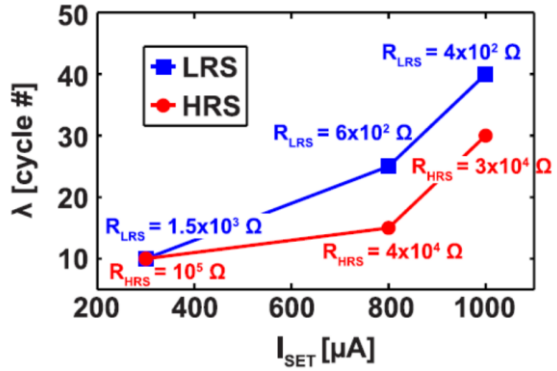


Figure 3 λ increases with current compliance. From [5]

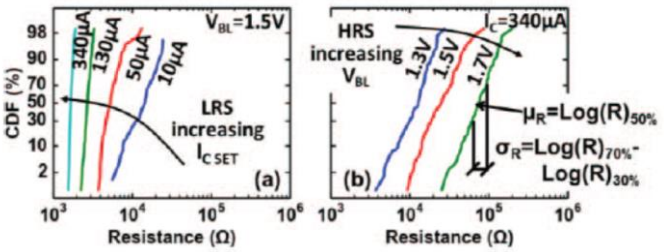


Figure 5 As the resistance increases, σ increases up to a saturation in HRS. From [6]

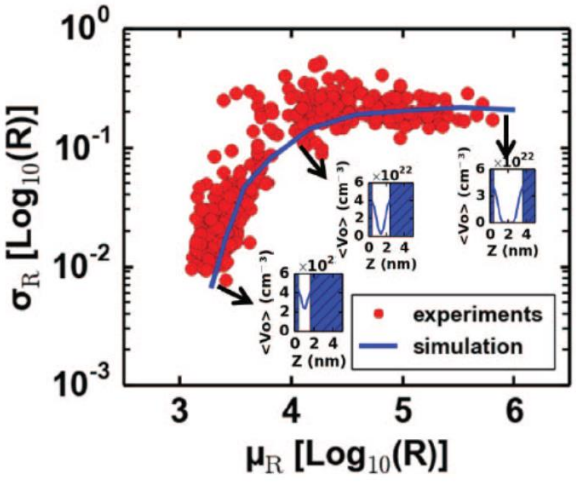


Figure 6 σ increases with the resistance. The insets refer to the model in Fig. 7. From [6]

II. Internal Correlation

When analyzing the C2C variability a memory effect in the short cycle range is in principle to be expected. This is due to the fact that the conductive filament (CF) breaks and reforms quite consistently in the first cycles, and one might expect that before the CF shape modifies enough, several write/erase cycles must be performed. This was studied and analyzed by *G. Piccolboni et al* in 2015 [5]. The short range correlation can be easily observed by looking at data in **Fig. 1**. The author shows the resistance cycling sampling the programmed LRS and HRS at three different times during a 1Mcycle test. Data show how consecutive HRS values are close to each other, while the average HRS modifies during the test. To quantify the correlation among consecutive cycles, the author showed the results of **Fig. 2**. On the left side, the color code represents the correlation coefficient (see [5] for details about the computation) of the resistance of a set of memory cells at a given cycle, say N, with the consecutive cycles N+1, N+2 and so on. Trivially, the main diagonal shows the autocorrelation always equal to 1. Going outside of the main diagonal reduces the correlation to lower values. The right side of **Fig. 2** shows a horizontal cut of the data set. This clarify how the correlation diminishes monotonically from values that are close to 1, down to much lower values. Eventually, after λ cycles, the correlation becomes constant at about 0.5-0.6, meaning that the resistance arrays are not correlated anymore at a distance higher than λ. This effect impacts on the dispersion of the resistance (in particular for HRS) when C2C and device-to-device (D2D) variability are considered at the same time. This is caused by the fact that the resistance values coming from the same array cycled just a few times will include correlated values. Eventually this will translate in a less dispersed resistance distribution. In **Fig. 3**, this is clarified by calculating the standard deviation of the difference between the resistances at cycle M, minus the resistances of a reference cycle 1. The σ increases linearly as the cycle difference is lower than λ and it stabilizes afterward. As a contrast, **Fig 3** shows the same calculation for an uncorrelated set of data. In this case, no σ increase is observed, as no short range correlation exists. This effect implies that when considering C2C variability it is mandatory to sample a cycle range wider than λ to fully estimate the resistance spread. On the other side, for very specific applications like neuromorphic circuit for image recognition, this effect can provide advantages. Since only a few cycles are required, the reduced variability can simplify the circuit structure and reduce the overall cost and complexity [5]. We show a last evidence in this section, concerning the link between the programmed resistance and the λ parameter. **Fig. 4** reports how λ increases as the resistance reduces. This is quite intuitive since a lower resistance implies a larger CF.

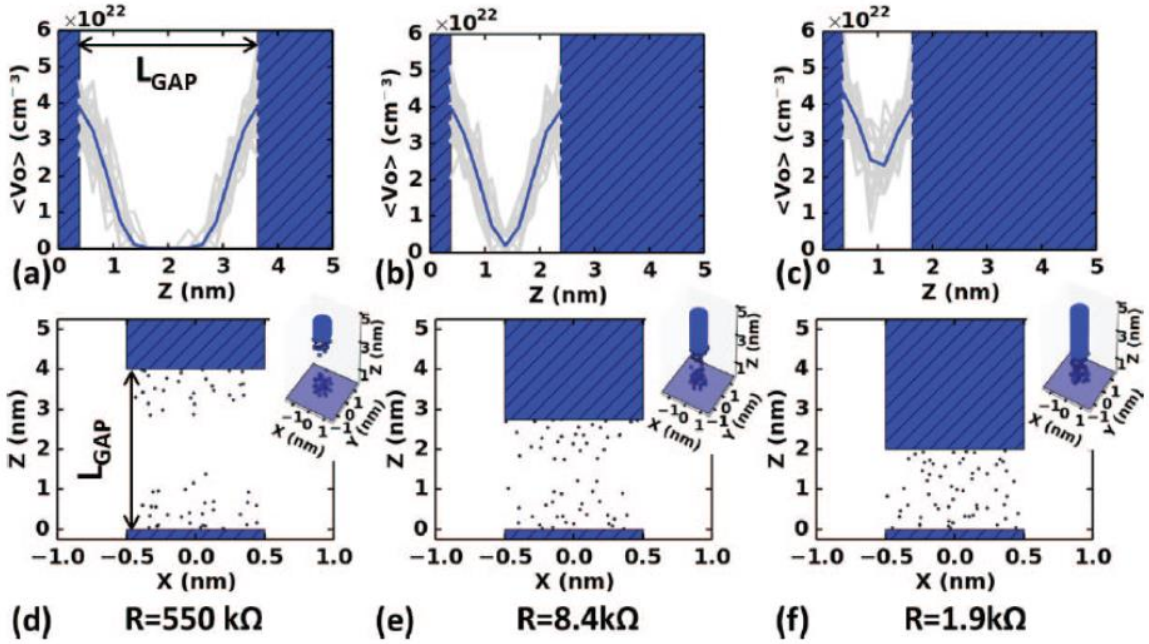


Figure 7 Modeling of the defect distribution within the gap. From (a) to (c) the defect density along the z-axis. From (d) to (f) a horizontal cut along the x-axis. From [6].

It is then not surprising that the correlation length is longer. In the next session, we further explore the link between the resistance and the variability.

III. Resistance impcat on variability

It is well established that both LRS and HRS distributions behave to a log-normal law. At the same time, it is a matter of fact that the LRS distribution, driven by the current compliance imposed by the selector device is much tighter than the HRS one. We are therefore induced to think that there must be a smooth transition between LRS and HRS. This is commonly observed in RRAM. In 2015, *D. Garbin et al.* investigated this transition and provided a simple model to capture the main physical aspects involved in that [6]. **Fig 5** reports some distributions of LRS and HRS programmed cells. These data show the increase of the resistance spread as the resistance is higher. The resistance level is controlled by the current compliance (I_{cc}) during set, while the off resistance can be modulated by the reset voltage. The author observed that as the off resistance continues increasing, the variability reaches a saturated value and does not evolve anymore [5]. This was captured by the model summarized in **Fig. 7**. The resistance is computed by a resistors network where each element has a resistance given by the electron tunneling probability among the defects that fill the gap. The defect density varies across the gap and its profile is strongly influenced by the gap length. As the gap shrinks, the defect density becomes almost constant. Clearly the resistance is linked to the defect distribution, the defect density and the gap length, while the variability is produced by the number of different defect distributions that can fit in the gap. As the CF widens a larger volume is available to accommodate multiple defects configurations. This produces an increase in the resistance relative spread (see **Fig. 6**). As the CF widens enough, a defect-free region forms in the middle of the gap, causing the saturation of σ showed in **Fig. 6** which corresponds to a fully HRS state.

Although the saturation of σ vs the resistance would be very beneficial for the applications, additional analysis showed that as a larger resistance range is considered, the variability tends to further increase. This was demonstrated for instance in [5] from which **Fig. 8** is reproduced. The picture shows that in the resistance range 1MΩ-1GΩ, the variability (here considered as the difference between the 70% and 30% of the resistance cumulative distribution) becomes even larger. This in turn, suggests that there is an intrinsic variability limit linked to the technology itself. This limit is explored in the next session, where more data, from a larger memory array, are considered.

IV. Variability limit

More recently a detailed study of the variability issues on HfO2-based RRAM arrays was presented [7]. This work addressed two main points. First it confirmed the behavior observed in Fig. 8 and compared this result with other studies showed on literature. This is displayed in Fig 9. The plot shows the σ vs median resistance (<R>) reported by several authors. As can be seen all the results are contained in a restricted region described by two boundary lines. This link between σ and <R> is captured assuming that the defects contributing to the conduction follow a Poisson distribution [8]. This gives the simple equation $\sigma = (G_0 \exp(\alpha\phi))^{1/2} \langle R \rangle^{1.5}$ (1), where G_0 is the quantum of conductance [7, 9], α and ϕ parameters that describe the parabolic potential barrier generated by the gap. This equation holds in the limit $R > 1/G_0$.

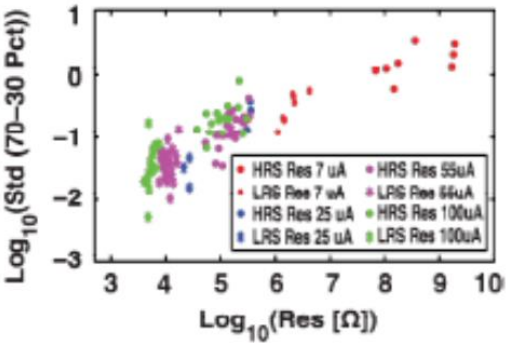


Figure 8 Variability vs Resistance. No saturation is observed up to 10^9 Ω. From [5]

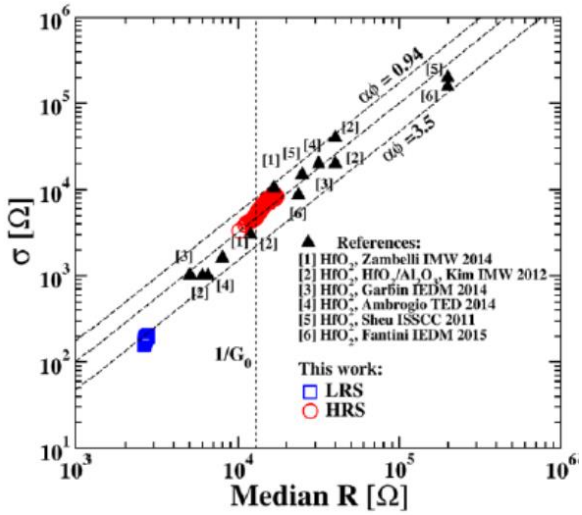


Figure 8 Comparison of σ vs R in different studies, showing that the same behavior is observed in HfO_2 RRAM. From [7]

Moreover, it was showed that for lower resistances, thus below $1/G_0$, the variability is allowed to reduce below the limit of **Eq. 1** and is proportional to σ_{Icc} $<R>^2$, where σ_{Icc} is the dispersion of the compliance current during set. The latter is linked to the selector technology and provides a tighter resistance distribution. These results show that RRAM resistance distributions are predictable and stable over cycling. Nevertheless as the reset resistances are not driven by any external limiting factor, their variability is lower-bound by an intrinsic spread which reflects the Poissonian behavior of the defect distribution. Eventually, we would mention a recent work by *A. Bricalli et al.*, reporting the σ vs <R> behavior of a different RRAM stack [4]. In this case a SiO_x-based RRAM is investigated. **Fig. 10** reproduces the data obtained with a comparison with HfO_x data. As can be seen SiO_x HRS variability does not follow anymore the same linear correspondence between σ and <R>. This indicates that different memory stack can break free from the limit we observed and described in **Eq. 1**. The reason for this needs to be investigated, however two pathways can be identified. First,

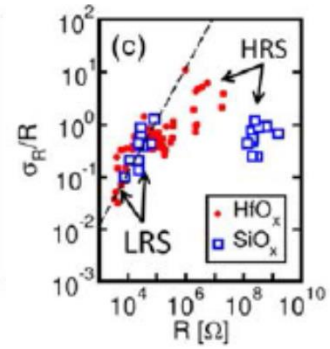


Figure 10 For SiO_x-based RRAM a lower σ was observed, suggesting that the variability limit can be lowered. From [4]

Eq. 1 could still hold but with a higher value of αϕ. Conversely a different switching mechanism could be involved, changing completely the hypothesis on the Poissonian defect distribution. *A. Bricalli* et al suggested that the carbon electrode used in the stack under investigation might have a major role in the resistance switching suggesting that the latter hypothesis could be correct.

V. Conclusion

HfO2-based RRAM shows a cycle-to-cycle correlation in a limited range of cycles. This range is a function of the programmed resistance. Beyond this limit, it is observed that an intrinsic link between s and the resistance exists. This limit prevents the HRS to be tighter than what **Eq. 1** predicts, while for LRS the limiting factor appears to be the selector device. Other material may show less restrictive limitation in HRS. SiO_x-based RRAM was recently shown to be possible example.

References

[1] Ueki et al. "Low-power ambedded ReRAM technology for IoT applications" International Conference VLSI, T108 2015.

[2] H. S. P. Wong et al."Metal-oxide RRAM", IEEE Proc, vol 100, no 6, pp. 1951-1970, 2012.

[3] E. Vianelloet al., "Resistive Memories for Ultra-Low-Power embedded computing design" Electron Devices Meeting (IEDM) IEEE International pp.6.3.1-6.3.4 15-17 Dec. 2014.

[4] A. Bricalli et al., "SiOx-based resistive switching memory (RRAM) for crossbar storage/select elements with high on/off ratio" Electron Devices Meeting (IEDM) IEEE International 3-7 Dec. 2016. in press.

[5] G. Piccolboni et al., "Investigation of the potentialities of Vertical Resistive RAM (VRRAM) for neuromorphic applications" Electron Devices Meeting (IEDM) IEEE Intern. pp.17.2.1-17.2.4, 7-9 Dec. 2015.

[6] D. Garbin et al.,"Modeling of OxRAM Variability from Low to High Resistance State using a Stochastic Trap Assisted Tunneling-based Resistor Network" International Conference ULIS, pp.125-129, 26-28 January 2015.

[7] A. Grossi et al., "Fondamental Variability Limits of Filament-based RRAM" Electron Devices Meeting (IEDM) IEEE International 3-7 Dec. 2016. in press.

[8] Perniola et al., "Universal Signatures from Non-Universal Memories: Clues for the Future…" International Memeory Workshop IMW IEEE International pp 1-3 15-18 May 2016.

[9] X. Lian et al. « Quantum point contact model of filamentary conduction in resistive switching memories » International Conference ULIS, pp.101-104, 6-7 March 2012.

ReRAM ON/OFF Resistance Ratio Degradation Due to Line Resistance Combined with Device Variability in 28nm FDSOI technology

H. Aziza*, P. Canet, J. Postel-Pellerin, M. Moreau, J.-M. Portal, M. Bocquet
Aix Marseille Univ., CNRS, IM2NP UMR 7334, 60 rue F. Joliot-Curie, 13453 Marseille Cedex 13, France
*hassen.aziza@im2np.fr

***Abstract*—Common problems with oxide-based Resistive RAM (ReRAM) are related to high variability in operating conditions and low yield. At array level, ReRAM memory cells suffer from different voltage drops seen across the cells due to the line resistances. Although research has taken steps to resolve these issues, variability combined with resistive paths remain an important characteristic for ReRAM. In this paper, the performance and reliability of ReRAM memory arrays is investigated in a 28nm FDSOI technology versus interconnects resistivity combined with device variability.**

Keywords- Resistive RAM; oxide-based RAM; variability; voltage drop (IR drop); reliability, 28nm FDSOI technology.

I. INTRODUCTION

At sub 32nm node, size reduction increases the resistivity of interconnects, inducing a voltage drop along the memory matrix lines, which can cause reliability issues. For new emerging memory technologies such as Resistive RAMs (ReRAMs), the data is stored as two resistance states of the resistive switching device [1], which makes these memories sensitive to resistive paths [2]. On top of that, common problems with ReRAMs are related to high variability in operating conditions [3].

II. RERAM OPERATION

In this study, we consider a 1T-1R ReRAM cell (one MOS transistor in series with one resistor) placed in a 32-bit word presented in Fig. 1. The word selection is achieved through the Word Line (WL) before any operation. Word programming is performed in 2 steps, considering that the FORMING operation occurs only once in the product lifetime. Once the word is selected, all cells are RESET in parallel (logical "0") through the Reset Word Line (WL_R), then memory cells are all SET (logical "1") through the Bit Line (BL). The WL_R line drives the whole word RESET current, making it sensitive to line resistivity.

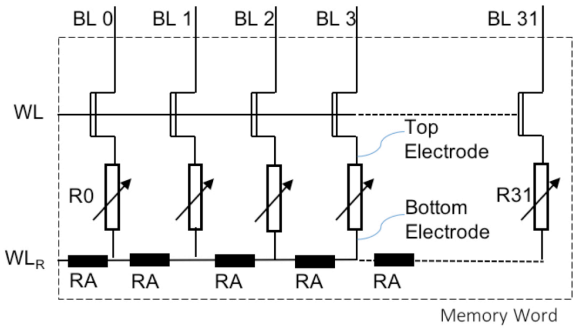


Figure 1. 32-bit ReRAM memory word

Indeed, during a memory word program operation, a voltage drop occurs along the WL_R line, which can be critical in terms of programming efficiency [2]. To monitor the programming efficiency, High Resistance State (R_{HRS}) and Low Resistance State (R_{LRS}) resistances are extracted after RESET and SET operations respectively.

III. SIMULATION AND RESULTS

A. Simulation

The memory cell is simulated with a bipolar ReRAM model calibrated on silicon [4] for a 28nm FDSOI technology. The ReRAM model features a variability dependency extracted on experimental data as presented in Fig. 2a (I-V characteristics in logarithmic scale). Typical values for LRS and HRS are R_{LRS}=10kΩ and R_{HRS}=130kΩ for the considered technology. The WL_R line resistance per cell (including vias) is called R_A (Fig. 1) and is evaluated to 1.6Ω [2].

B. Results

Impact of line resistances on R_{HRS} and R_{LRS} is presented in Fig. 2b. The voltage drop induced by line resistances results in a weak FORMING (i.e. high resistive path between ReRAM electrodes). SET and RESET programming levels are thus impacted resulting in a higher R_{HRS} (+5%) and R_{LRS} (+30%) values for cells located at the end of the memory word [5]. Impact of

line resistances combined with variability is presented in Fig. 3.

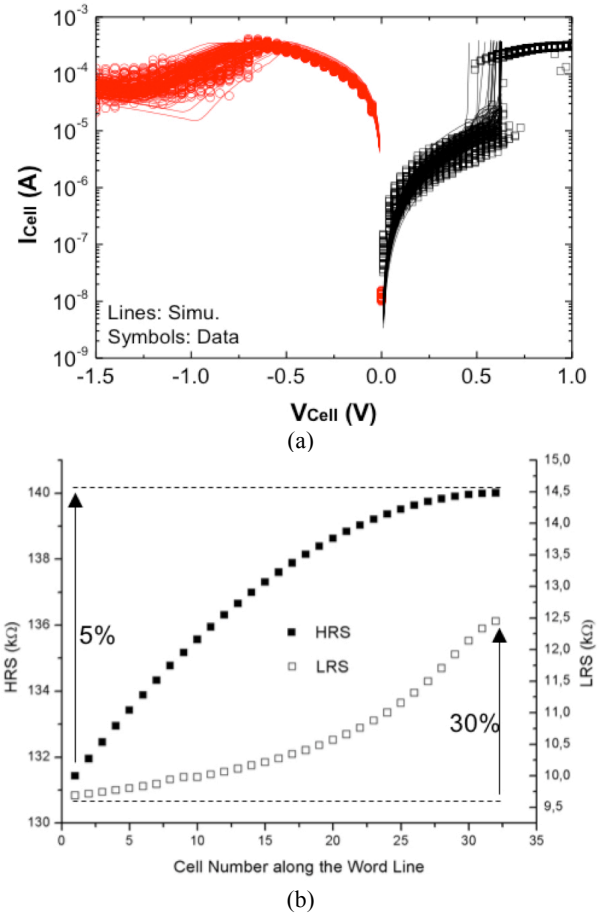


Figure 2. (a) Measured and corresponding simulated I-V characteristic obtained from TiN/Ti/HfO2/TiN devices showing strong variation on R_{HRS} and R_{LRS}
(b) Resistances (HRS & LRS) along the WL

After 300 Monte Carlo runs, R_{HRS} and R_{LRS} distributions are extracted for 1 cell over 4 and displayed in box-plot forms. Fig. 3a shows that LRS distribution spread increases along the word line. Besides, anomalous LRS residual cell populations (far from their typical LRS values) are visible from cell 7. This effect is related to cells not properly formed due to line resistances. In Fig.3b, HRS distributions are presented. One can notice the HRS mean distribution shift to higher values with the word length. For comparison purposes, Fig. 4 presents the full range variation (from min to max) of each box-plot in HRS and LRS. Along the WL, HRS/LRS distributions are getting closer and closer due to marginal LRS cells, affecting memory reliability.

IV. CONCLUSION

Variability and voltage drop due to resistive paths present a major challenge for fabrication process and design engineers in ReRAM memories. In the proposed study, impact of variability combined with IR drop is evaluated based on a 32bits 1T-1R ReRAM memory word. HRS/LRS resistance ratio is extracted for each memory cell to assess the technology reliability. The study provides a valuable feedback for designers during Resistive RAM memory array sizing.

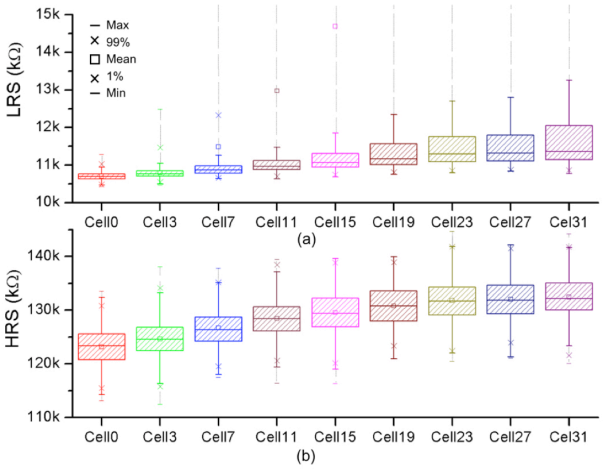


Figure 3. (a) LRS and (b) HRS distributions along the WL

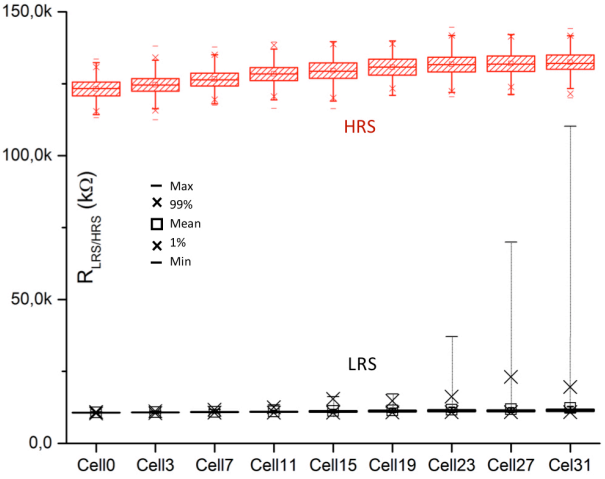


Figure 4. LRS and HRS distributions overlaps along the WL

REFERENCES

[1] S. Hamdioui et al. "Memristor based memories: Technology, design and test", in Design & Technology of Integrated Systems In Nanoscale Era Conference, pp. 1.7, 2014.

[2] Liang & al., "Effect of wordline/bitline scaling on the performance, energy consumption, and reliability of cross-point memory array", ACM Journal on Emerging Technologies in Computing Systems, Vol. 9, No. 1, 2013.

[3] E.I. Vatajelu et al., "Nonvolatile memories: Present and future challenges", Design & Test Symposium (IDT), pp. 61-66, 2014.

[4] M. Bocquet et al., "Robust compact model for bipolar oxide-based resistive switching memories", IEEE Transactions on Electron Devices, (61), 674-681, 2014.

[5] P. Canet et al., "Impact of Resistive Paths on NVM Array Reliability: Application to Flash & ReRAM", Microelectronics Reliability, (64), 36-41, 2016.

Static Noise Margin Analysis of 8T TFET SRAM Cells Using a 2D Compact Model Adapted to Measurement Data of Fabricated TFET Devices

F. Horst[1,2,✠], M. Graef[1,2], F. Hosenfeld[1,2], A. Farokhnejad[1,2], G. V. Luong[3], Q. T. Zhao[3], B. Iñíguez[2], A. Kloes[1]
[1]NanoP, TH Mittelhessen University of Applied Sciences, Giessen, Germany,
[2]DEEEA, Universitat Rovira i Virgili, Tarragona, Spain, [3]PGI 9-IT and JARA-FIT, Forschungszentrum Jülich, Germany
✠fabian.horst@ei.thm.de

***Abstract*—In this paper a static noise margin (SNM) analysis is done for a 8T SRAM cell built up with complementary tunnel-FETs (TFETs). Simulations are done with the help of a Verilog-A implementation of a 2D DC compact model of a double-gate (DG) TFET, published in [1]. The compact model has been adapted to fabricated nanowire (NW) GAA TFETs before analyzing the SNM of the 8T SRAM cell. The impact of the ambipolar behavior and the unidirectional current of TFETs on the SRAM cell simulation are considered and analyzed in this work.**

***Keywords*—TFET, SRAM, Compact Model, Double-Gate (DG), Static Noise Margin, Nanowire GAA TFET, Circuit Simulation.**

I. Introduction

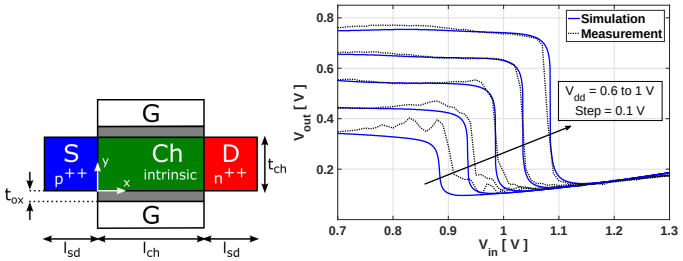
The persistent aim to reduce the supply voltage in circuit design is one of today's biggest challenges. The conventional CMOS technology has reached its physical limitations for ultra-low power applications. A promising candidate to replace this technology is the TFET because of its possibility to obtain subthreshold slopes below 60 mV/dec and small leakage currents [2]. But there is still room for improvement, regarding the ambipolar behavior or the unidirectional current of the TFETs which plays an important role in the circuit design and simulation. In [3], [4], these attributes are taken into account in the SRAM cell design and the subsequent analyses, although these papers are based on time consuming TCAD simulations. This work presents a SNM analysis of a 8T TFET SRAM cell using a 2D DC compact DG TFET model, which is adapted to the fabricated GAA devices [1].

II. Compact Modeling Approach

This section presents a short review of a 2D physics-based compact model for a DG TFET (see Fig. 1(a)), which was introduced in [1]. Following, the modeling approach for the on-state of the device is explained, which is transfered to the ambipolar-state at the end of this section.

Firstly, the electrostatic potential along the x-axis for an arbitrary y-position is approximated at the source-to-channel junction with the help of mathematic functions using an analytical closed-form potential solution as presented in [5]. Applying the compact potential solution, one is able to describe the band structure of the device. Next, the tunneling generation rate $TGR(x, y)$ is approximated along the x-axis for any y-position with the help of a Gaussian distribution using the tunneling probability, the Landauer transmission theory as well as the electric field. The tunneling probability is calculated by the Wentzel-Kramers-Brillion approximation in dependency of the band structure. Integrating the TGR leads to the current density along the y-axis $J_y(y)$, whereby J_y must be described in a compact form as well. This description is also done by a Gaussian distribution, which leads to the compact current density $J_{comp}(y)$ term. The subsequent integration along the y-axis results in the compact tunneling current expression for the on-state of the device $I_{comp,on}$. The compact ambipolar current expression $I_{comp,amb}$ is calculated in the same way as $I_{comp,on}$ considering the drain-to-channel junction and the physical parameters of holes. A summation of both parts leads to the total compact tunneling current term of the device [1].

By using the compact model, it is possible to simulate the voltage transfer curve (VTC) of a fabricated single stage TFET inverter [6] for different supply voltages V_{dd} [1], shown in Fig. 1(b). The ambipolar behavior of the TFETs causes an output voltage degradation of the inverter VTC, which has a negative impact on the SRAM behavior.



(a) (b)

Figure 1. (a) 2D device geometry of the n-type DG TFET compact model [1]. (b) Simulation results of the complementary TFET inverter for different supply voltages V_{dd} compared to measurement data [1].

III. 8T TFET SRAM Cell Layout

The simulated SRAM cell layout consists of 8 complementary TFET devices (Fig. 2). The conventional 6T SRAM layout cannot be used in this work due to the unidirectional device current of TFETs [3], [4]. The transistors $T_1 \dots T_4$ describe the two cross-coupled inverters, whereby T_2 & T_4 are the pull-up (PU) and T_1 & T_3 are the pull-down (PD) transistors. T_5 & T_6 are the outward-facing access (AT) transistors for the write operation. The read transistors (RT) T_7 & T_8 form a decoupled read circuit. The word lines for read/write operation are denoted as $\mathrm{WL_{R/W}}$, where the bit lines for write operation are marked with $\mathrm{BL(B)_W}$ and the read bit line with $\mathrm{BL_R}$.

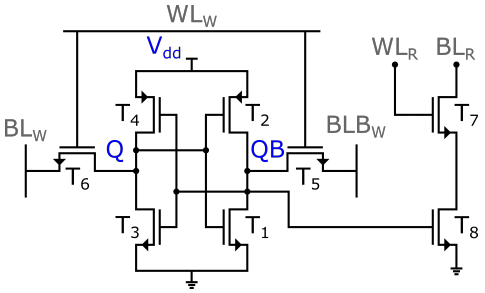
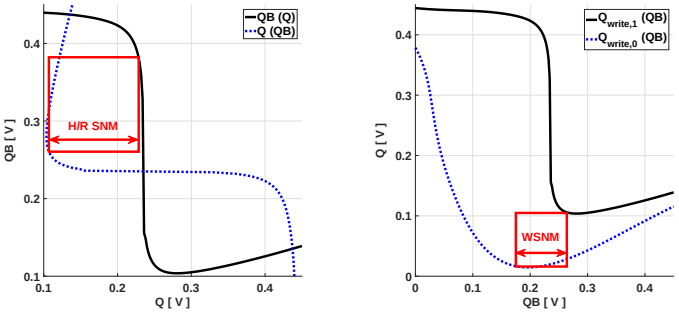


Figure 2. Schematic 8T TFET SRAM cell layout [3].

IV. Simulation Results

The simulation of the 8T SRAM cell is done using the compact model, presented in section II and the device modeling software IC-CAP from Keysight Technologies. In order to obtain simulation results close to reality, the compact model is adapted to fabricated complementary NW GAA TFETs and used to simulate fabricated TFET inverters (Fig. 1(b)) [1], [6].

Firstly, the butterfly curve for a hold/read (H/R) operation is simulated to get the resulting H/R SNM for $V_{dd} = 0.7\,V$. For simulating a hold operation, all bit ($BL(B)_X$) and word lines (WL_X) stay at zero volts. For a read operation, BL_R is pre-charged to $V_{dd}/2$ and WL_R is set to V_{dd}. The simulation results are shown in Fig. 3(a), whereby the related SNM is calculated as in [7]. The resulting butterfly curves for hold and read are identical, which is caused by the decoupled read circuit (Fig. 2). In Fig. 3(b) the simulation result for a write (W) operation is shown. To write a logical '1', the bit lines BL_W and BLB_W are pre-charged to $V_{dd}/2$, whereby WL_W is set to V_{dd}. To write a '0', only WL_W is adjusted to V_{dd}, where BL_W and BLB_W stay at zero volts. The WSNM is calculated in the same way as mentioned above.

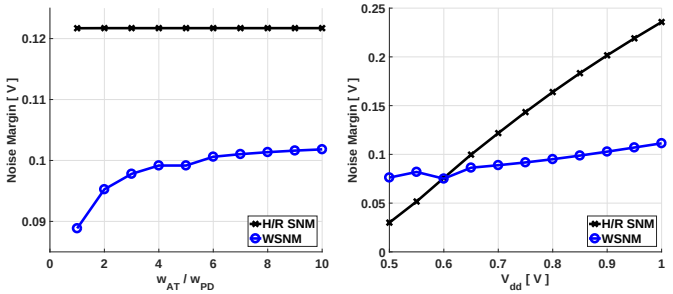


(a) H/R SNM (b) WSNM

Figure 3. (a) Butterfly curve of the 8T SRAM cell for hold/read operation to analyze the H/R SNM, simulated for $V_{dd} = 0.7\,V$. (b) Inverter curves for different write operations ('0'&'1') to analyze the WSNM, simulated for $V_{dd} = 0.7\,V$. These curves are simulated for $w_{PD} = w_{PU} = w_{AT} = w_{RT}$.

The second simulation is done for various device widths of the AT (w_{AT}) in order to show the influence on the SNM. As it can be seen in Fig. 4(a), the H/R SNM stays constant for an increasing w_{AT} due to the decoupled read circuit. Furthermore, the WSNM increases for an increasing w_{AT}. Because a wider AT causes a better pull-down behavior of the inverters.

Fig. 4(b) shows the resulted SNM for various V_{dd}. The increasing V_{dd} causes a bigger on/off ratio of the inverter (see Fig. 1(b)), hence the H/R SNM increases. Moreover, it can be seen that a smaller V_{dd} causes a smaller WSNM. The reasons for that are the unidirectional current of the AT, which deprives the push-pull action during the write operation [4]. The ambipolar behavior of the TFETs causes an output voltage degradation of the inverter VTC (see Fig. 1(b)).



(a) (b)

Figure 4. (a) Simulation results of H/R & WSNM for various w_{AT}, whereby $w_{PD} = w_{PU}$. (b) Simulation results of H/R & WSNM for various V_{dd}, with $w_{PD} = w_{PU} = w_{AT} = w_{RT}$.

V. Conclusion

This work presents a static noise margin analysis of a 8T TFET SRAM cell. The simulations are done with the help of a 2D DC compact model for DG TFETs, whereby the model has been adapted to fabricated complementary NW GAA TFETs and fabricated TFET inverters. The simulations show the impact of various widths of the access transistors and various V_{dd} on the resulting hold/read and write SNM.

Acknowledgment

This project was supported by the German Federal Ministry of Education and Research under contract No. FKZ 03FH010IX5 and the Spanish Ministry of Economy and Competitiveness through project GREENSENSE (TEC2015-67883-R). We would like to thank Keysight Technologies for the license donation and support of the software IC-CAP and AdMOS GmbH for support.

References

[1] F. Horst *et al.*, "Implementation of a DC Compact Model for Double-Gate Tunnel-FET Based on 2D Calculations and Application in Circuit Simulation," in *Proc. 46th European Solid-State Device Research Conf. (ESSDERC)*, Sep. 2016, pp. 456–459.

[2] H. Lu and A. Seabaugh, "Tunnel Field-Effect Transistors: State-of-the-Art," *IEEE J. Electron Devices Soc.*, vol. 2, no. 4, pp. 44–49, 2014.

[3] S. Strangio *et al.*, "Impact of TFET Unidirectionality and Ambipolarity on the Performance of 6T SRAM Cells," *IEEE Journal of the Electron Devices Society*, vol. 3, no. 3, pp. 223–232, May 2015.

[4] Y. N. Chen *et al.*, "Evaluation of Stability, Performance of Ultra-Low Voltage MOSFET, TFET, and Mixed TFET-MOSFET SRAM Cell With Write-Assist Circuits," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 4, no. 4, pp. 389–399, Dec. 2014.

[5] M. Graef, T. Holtij, F. Hain, A. Kloes, and B. Iñíguez, "A 2D Closed Form Model for the Electrostatics in Hetero-Junction Double-Gate Tunnel-FETs for Calculation of Band-to-Band Tunneling Current," *Microelectronics Journal*, vol. 45, no. 9, pp. 1144–1153, 2014.

[6] G. V. Luong *et al.*, "Experimental Demonstration of Strained Si Nanowire GAA n-TFETs and Inverter Operation With Complementary TFET Logic at Low Supply Voltages," *Solid-State Electronics*, vol. 115, pp. 152–159, 2016.

[7] E. Seevinck, F. J. List, and J. Lohstroh, "Static-Noise Margin Analysis of MOS SRAM cells," *IEEE Journal of Solid-State Circuits*, vol. 22, no. 5, pp. 748–754, Oct. 1987.

Design and Implementation Methodology of Energy-Efficient Standard Cell Memory with Optimized Body-Bias Separation in Silicon-on-Thin-BOX

Yusuke YOSHIDA and Kimiyoshi USAMI
Shibaura Institute of Technology, 3-7-5 Toyosu, Koto-ku, Tokyo, Japan

***Abstract*—This paper describes a design of energy-efficient Standard Cell Memory (SCM) using Silicon-on-Thin-BOX (SOTB). We present automatic place and routing (P&R) methodology for optimal body-bias separation (BBS) for SCM to reduce leakage current without major sacrifice of performance. Simulation results demonstrated that proposed automatic P&R methodology can reduce wire length by 22% and provide more than 50% smaller energy consumption reduction than conventional SCM of implementing by the standard digital design flow at Near/Sub-Vth region operation.**

I. INTRODUCTION

Standard Cell Memory (SCM) has been drawing attention as a technique to realize an embedded memory macro operating at near-Vth or sub-Vth (Near/Sub-Vth) region [1]. Although SRAMs with more than 6 transistors [2] for a bit-cell operate stably at such ultra-low voltage, they suffer from large silicon footprint, large leakage current and unavailability as a standard memory macros provided by library venders. In contrast, SCM can be synthesized, placed and routed with a standard cell library including latch, multiplexor (MUX) and logic cells. In [3], the authors demonstrated that a 1Kbyte SCM operates correctly at the supply voltage as low as 0.3V and achieves smaller read/write power dissipation with shorter access time than its SRAM counterpart. Paper [4] presented that energy of SCM can be reduced by improving the control scheme of the MUX tree in the read-out circuit. However, fine-grain body-biasing techniques for SCM to reduce leakage current have not been discussed in these papers. We proposed an idea to give different body biases to latch cells to store data and MUX cells in the read-out circuit [5], but have not proposed an automated implementation flow for SCM to optimize energy while considering different constraints at read and write timings.

In this paper, we describe an SCM structure to realize optimal body-bias separation (BBS) on a Silicon-on-Thin-BOX (SOTB) device [6] and propose an automatic place and route (P&R) technique to implement the entire SCM employing multiple body-bias voltages. We also present experimental results showing that our approach produces SCM macros with higher Fmax and lower dynamic/ leakage energy compared to the conventional SCM.

II. SCM ARCHITECTURE

Figure 1 shows the structure of our SCM. The SCM consists of a Controller module and LatchMUX module. The Controller module includes an address decoder and read/write control circuits, whereas the LatchMUX module consists of latch cells to store data and MUX cells in the read-out circuit for one column.

In this paper, we assume that SCM performs the read operation within one clock cycle and does the write operation as well. For address decoding and read/write timings, we put the following assumptions:

- Write operation
  The write operation is done by converting one-hot address at the write address decoder (WAD) in the first half clock cycle (referred to as WAD process) and by providing the write signal to latch cells based on the one-hot address in the second half clock cycle (referred to as LatchWrite process).
- Read operation
  The read operation is done by converting one-hot address at the read address decoder (RAD) and selecting the output word in MUX in one clock cycle.

Above timing constraints lead to three critical paths (i.e. half cycle for WAD, half cycle for LatchWrite process and one cycle for read operation).

III. PROPOSED AUTOMATIC P&R METHODOLOGY

In this section, we describe a methodology for applying body bias separation (BBS) between latch cells and MUX cells to reduce leakage current without major sacrifice of performance. Further, we propose an implementation methodology to realize BBS using automatic P&R tool and to optimize timing constraints and placement constraints.

Due to the design rule for the inter-well spacing in the triple-well structure and resulting area overhead, reverse-body-bias (RBB) control is applied at every two row for nMOS (nRBB) and is applied to the entire cells for pMOS (pRBB). Moreover, it turned out that the latch is on a non-critical path except LatchWrite process which has enough slack. As a result, SCM can reduce leakage current without major sacrifice in performance by applying strong nRBB to latch cells and weak nRBB to MUX cells separately. This approach requires regularly structured layout with separate rows for latch cells and MUX cells. Figure 2 shows the floorplan for the proposed layout to realize BBS. This floorplan also enables to reduce wire length because most wires are completed within the module and routed straight from Controller module to LatchMUX module.

To deal with half-clock based timing constraints for the write operation and one-clock based constraints for the read operation, we divide the layout process into two steps:

First, the timing of write paths are optimized by applying automatic P&R using only Controller module and providing the timing constraint of half clock cycle. Second, the timing of read path is optimized by providing the timing constraint of one clock cycle and applying automatic P&R using the entire SCM including the first layout in which “dont_touch” is partly applied and latch cells are excluded from the clock tree synthesis.

IV. SIMULATION RESULT

We conducted HSPICE simulations with the parasitic R and C extracted from the layout for our proposed SCM

(referred to as BBS) generated by our P&R flow, and for the conventional SCM (Conv-SCM: no body bias separation) created by the standard digital design flow.

Both of these SCMs were designed in commercial 65nm SOTB technology. The memory size is 4Kbit (128word×32bit) and the timing constraints of P&R are the same at both memories.

A. Fmax

Figure 3 shows the maximum frequency (Fmax) of BBS and Conv-SCM. Fmax of BBS is 1.75X higher than Conv-SCM. It was found that in Conv-SCM the write operation becomes a bottleneck at Fmax.

B. Energy Consumption

Figure 3 and 4 show energy consumption per bit-access with zero-body-bias (ZBB) and RBB. It is revealed that BBS can achieve 23% smaller read dynamic energy and 55% smaller write dynamic energy than Conv-SCM with ZBB. This is because BBS can reduce the wire length by 22% and reduce the size of the clock tree for the write operation.

Figure 4 shows that BBS using RBB provides more than 50% smaller energy consumption than Conv-SCM using RBB at Near/Sub-Vth region. This is because BBS can reduce leakage energy consumption which heavily influences the total energy consumption at Near/Sub-Vth region. BBS using RBB is particularly energy efficient at Near/Sub-Vth region because this approach can reduce leakage energy without major increase of delay. We also found that the proposed SCM operates at the minimum energy point (0.3V) with 3.79fJ energy per bit-access and outperforms the reported Sub-Vth memories as summarized in TABLE 1. The area of proposed SCM is 5% smaller than Conv-SCM despite realizing BBS and approximately 2 times larger than that of Sub-Vth SRAM [2].

C. Variation Tolerance

Figure 5 shows leakage energy at 0.3V under the temperature and process variations when applying the adaptive body bias (ABB) technique. We evaluated leakage energy consumption of BBS and Conv-SCM under the temperature and process variations when ABB was applied for the target frequency of 15 MHz. This frequency is Fmax of Conv-SCM under TT and 25ºC. Not only correction of Vth by using ABB of BBS can operate the same performance as TT at every temperature, but also it can reduce leakage energy consumption by approximately 70% compared to Conv-SCM using ABB at every temperature.

REFERENCES

[1] P. Meinerzhagen, et al., ESSDERC, Sep. 2012.
[2] S. Lutkemeier, et al, IEEE J. Solid-State Circuits, Jan. 2013.
[3] A. Teman, et al, ACM TODAES Vol.21, Sep. 2016.
[4] J. Shiomi, et al, ISQED, Mar. 2015.
[5] Y. Yoshida, et al, ITC-CSCC, Jul. 2016.
[6] R. Tsuchiya, IEDM, Dec. 2004.

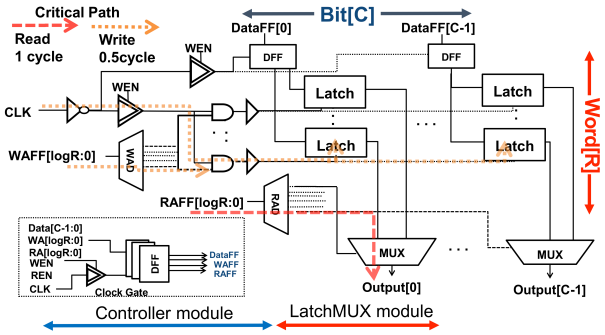


Figure 1. Block diagram of proposed SCM

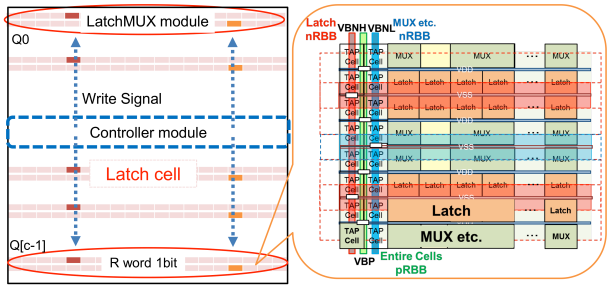


Figure 2. Floorplan of module and cells

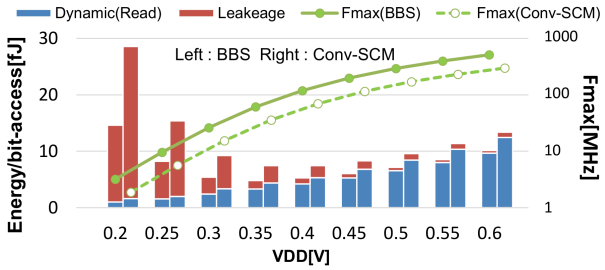


Figure 3. Fmax and energy consumption with ZBB

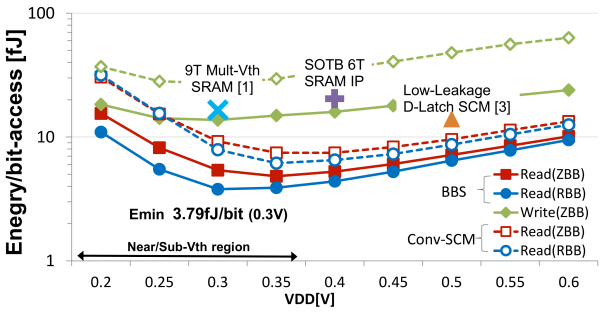


Figure 4. Energy consumption of read and write operation

TABLE I. COMPARISON WITH PRIOR SUB-VT MEMORIES IN 65NM

	SRAM	SCM		
	Multi-Vth 9T [2]	*Low-Leak Latch [1]*	*Conventional*	***This Work***
Memory capacity	2Kbit	4Kbit	4Kbit	**4Kbit**
VDDmin[mV]	280	420	200	**200**
Emin/bit[fJ/bit]	16.5	14	6.15(0.35V)	**3.79**(0.3V)
Fmax [MHz]	0.5(0.3V)[a]	0.11(0.5V)[a]	10(0.35V)[a]	**6.2**(0.3V)[a] **0.7**(0.3V)[b]
Area/bit[μm²]	4.61	12.7	10.7	**10.2**
Leakage/bit[pW]	17.8(0.3V)	0.5	145(0.35V)[a]	**74**(0.3V)[a] **4.9**(0.3V)[b]

a. Minimum Energy Point b. Ultra-Low-Leak mode (at operation)

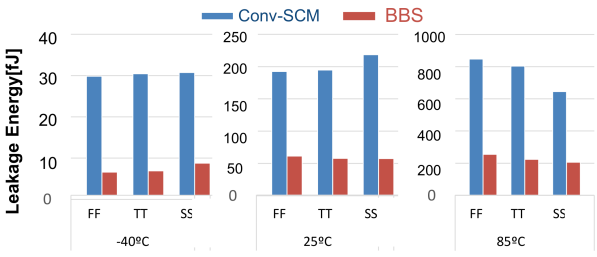


Figure 5. Leakage energy consumption at 0.3V under the temperature and process variations with ABB

Indium-Oxide Nanoparticles for Ox-RRAM in CMOS back-end-off-line

Edgar A. A. León Pérez, Pierre-Vincent Guenery, Oumaïma Abouzaid, Khaled Ayadi, Nicolas Baboux, Liviu Militaru, Abdelkader Souifi
Institute of Nanotechnologies of Lyon UMR CNRS 5270, INSA de Lyon, 69621 Villeurbanne Cedex, France

Jérémy Moeyaert, Thierry Baron
Laboratoire des Technologies de la Microélectronique, Centre National de la Recherche Scientifique, Grenoble, France

***Abstract*—We report on the fabrication and characterization of Resistive Random Access Memory (RRAM) devices based on nanoparticles with a memory capacitor structure. Our approach is based on the use of indium oxide (In_2O_3) nanoparticles (or nanocrystals NCs) embedded in a dielectric matrix as a charge trapping layer using CMOS-full-compatible fabrication processes in view of back-end-off-line integration for non-volatile memory (NVM) applications. Current-voltage characteristics (I-V) showed bipolar switching behavior for all the fabricated devices, with I_{ON}/I_{OFF} ratios up to 10^5. Moreover, our best structure yields up to 24 write/erase cycles, proving that our results provide insights for further integration of In_2O_3 nanoparticles-based devices.**

Keywords- RRAM; non-volatile memory; nanoparticles; indium oxide.

I. INTRODUCTION

Among new emergent memory technologies, such as Magnetic RAM (MRAM) or Conductive Bridge RAM (CBRAM), RRAM technology shows up as a suitable candidate for the next generations of non-volatile memories due to the commutation between two resistive states when thin layers of metal oxides are used in metal-insulator-metal (MIM) structures (two terminal devices) [1]–[5]. Furthermore, it has been reported that the integration of nanoparticles in these devices might directly improve their scalability and performance [6]–[14].

In this work, we show In_2O_3 NCs-based RRAM devices potential for CMOS back-end-off-line integration in future NVM technology. As a matter of fact, In_2O_3 is a suitable material due to its work function value (~5 eV) [15] and band-offset with SiO_2, which directly accounts for the charge retention on memory capacitor-like structures. We propose an insulating-NCs-insulating system which benefits from the above-mentioned characteristics for data retention. The fabricated RRAM structures $Au/Al_2O_3/In_2O_3$-NCs/SiO_2/Si-n+ (see Fig. 1 (a)) exhibited a bipolar switching behavior for different sizes of the devices (circular-shaped electrodes) going from D=500 µm down to D=50 µm.

II. EXPERIMENTS

In order to accomplish the above-mentioned-two-terminal-device architecture, n+-type silicon substrates were used. The first step of the process is the thermal oxidation of the substrate, in order to obtain a 2-nm-thick high quality SiO_2 layer (control oxide). A very thin layer of In is then deposited onto the SiO_2 by metal-organic-chemical-vapor-deposition (MOCVD), which is annealed (<450 °C) in order to obtain nano-sized droplets at its surface. The system is then cooled down to room temperature and exposed to ambient atmosphere, forming In_2O_3 NCs on the SiO_2 surface. Fig. 1 (b) shows an atomic force microscope (AFM) scan of these NCs with diameters between 9 nm and 14 nm, and an estimated density of $3x10^8$ NCs/cm^2. Subsequently, a 2 nm-3 nm-thick Al_2O_3 tunnel-oxide is deposited by atomic layer deposition (ALD) at 200 °C. In order to define the devices-dimensions, an UV-photolithography and evaporation processes were used to obtain circular-shaped gold electrodes (200 nm-thick) ranging from 50 µm to 500 µm. Finally, an annealing process of 10 minutes, at 400 °C under N_2 atmosphere, was used to passivate the electrical contacts.

Current-voltage measurements were then recorded using a Keithley 4200 semi-conductor parameter analyzer at room temperature (300 K).

III. RESULTS AND DISCUSSIONS

Bipolar switching behavior was observed for all the devices by sweeping the DC bias amplitude between ±2.5 V. On one hand, Fig.2 (a) illustrates the I-V characteristic for a 500 µm-diameter device. When the bias voltage was swept from 0 V to 2.5 V, the resistance of the device showed an initial high-resistance state (HRS) up to 1.2V, and then a gradual transition from HRS to the low resistance state (LRS) until 2.5 V (transition 1 → 2). When a negative sweep voltage was applied from 0 V to -2 V, the LRS state was maintained until -1 V, then the transition from LRS to HRS was observed (transition 3 → 4). This process was repeated up to 24 cycles (set-reset process).

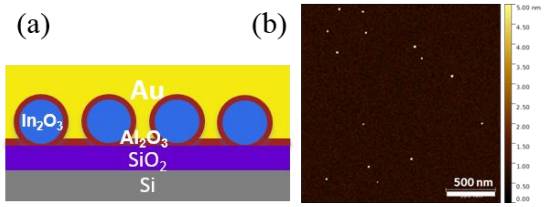


Figure 1. (a) Schematic cross-section structure of the fabricated devices, and (b) AFM cartography of one device after In_2O_3 NCs formation on top of SiO_2 surface.

It is worth noting that no initialization process (or "forming process") was needed to switch from HRS to LRS, only the sweeping voltage was sufficient. In this case, I_{ON}/I_{OFF} ratios of about 10^3 were observed for a read voltage (V_{read}) of 0.5V. On the other hand, for the 50 µm-diameter devices, besides the bipolar switch behavior, an initialization process was required to pass from the initial HRS state to LRS state. This process consisted of applying a 2.5 V bias pulse during 1.5s, after what the device was in a LRS state. Afterwards, lower bias amplitude were used in order to observe the switching behavior (from -1.2 V to 1V). The blue curve on the bottom in Fig. 2 (b) shows the behavior prior to the initialization process, showing no change while applying the bias sweep for positive and negative values. In contrast, the red curve shows the values for I after the initialization process, and it can be seen that for positive values of the applied bias, the LRS state is observed, while for negative values a transition from LRS to HRS is seen around -0.5 V. The black and green curves show two more similar cycles. In this case, I_{ON}/I_{OFF} ratios of about 10^5 were observed for V_{read}= 0.2V.

IV. CONCLUSION

The present work aimed at bringing the proof-of-concept of two-terminal devices for CMOS back-end-off-line integration for NVM applications. We presented a functional material stack of In_2O_3-NCs based devices and their electrical characterization, showing that I_{ON}/I_{OFF} ratios in the range 10^3-10^5 can be obtained. It is shown, for the first time, the bipolar switching behavior of RRAM devices based on In_2O_3 nanoparticles. This results provide insights in the investigation for In_2O_3 NCs-based RRAM devices integration.

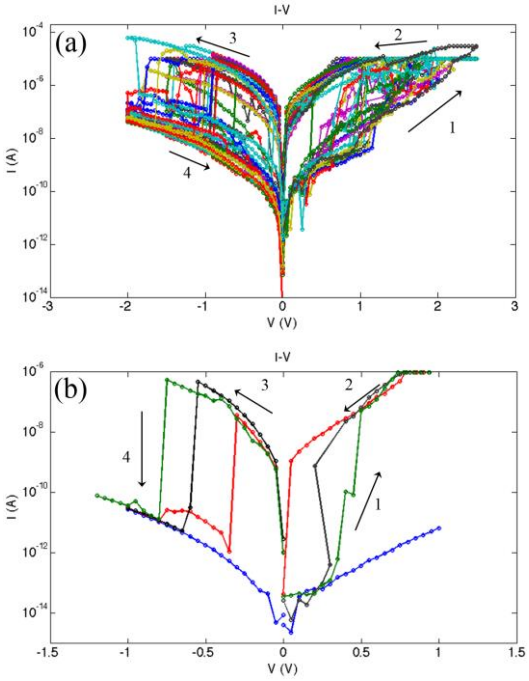


Figure 2. I-V characteristics of (a) a 500 µm-diameter device (24 cumulated cycles), and (b) a 50 µm-diameter device (3 cumulated cycles).

REFERENCES

[1] H. S. P. Wong *et al.*, "Metal – Oxide RRAM," *Proc. IEEE*, vol. 100, no. 6, pp. 1951–1970, 2012.

[2] A. Sawa, "Resistive switching in transition metal oxides," *Mater. Today*, vol. 11, no. 6, pp. 28–36, 2008.

[3] K.-C. Chang *et al.*, "Physical and chemical mechanisms in oxide-based resistance random access memory.," *Nanoscale Res. Lett.*, vol. 1, no. 10, p. 27p, 2015.

[4] F. M. Puglisi, D. Deleruyelle, J. M. Portal, P. Pavan, and L. Larcher, "A multi-scale methodology connecting device physics to compact models and circuit applications for OxRAM technology," *Phys. Status Solidi Appl. Mater. Sci.*, vol. 213, no. 2, pp. 289–301, 2016.

[5] X. Tian *et al.*, "Recent development of studies on the mechanism of resistive memories in several metal oxides," *Sci. China Physics, Mech. Astron.*, vol. 56, no. 12, pp. 2361–2369, 2013.

[6] T. C. Chang, F. Y. Jian, S. C. Chen, and Y. T. Tsai, "Developments in nanocrystal memory," *Mater. Today*, vol. 14, no. 12, pp. 608–615, 2011.

[7] J. a. Yater, "Implementation of Si nanocrystals in non-volatile memory devices," *Phys. Status Solidi*, vol. 210, no. 8, pp. 1505–1511, 2013.

[8] V. Kannan and J. K. Rhee, "Controllable switching ratio in quantum dot/metal-metal oxide nanostructure based non-volatile memory device," *Appl. Phys. A Mater. Sci. Process.*, vol. 108, no. 1, pp. 59–63, 2012.

[9] J. De Blauwe, "Nanocrystal nonvolatile memory devices," *IEEE Trans. Nanotechnol.*, vol. 1, no. 1, pp. 72–77, 2002.

[10] C. H. Cheng and A. Chin, "Nano-crystallized titanium oxide resistive memory with uniform switching and long endurance," *Appl. Phys. A Mater. Sci. Process.*, vol. 111, no. 1, pp. 203–207, 2013.

[11] Y.-T. Tsai *et al.*, "Influence of Nanocrystals on Resistive Switching Characteristic in Binary Metal Oxides Memory Devices," *Electrochem. Solid-State Lett.*, vol. 14, no. 3, p. H135, 2011

[12] L. Masoero *et al.*, "Scalability of split-gate charge trap memories down to 20nm for low-power embedded memories," *Tech. Dig. - Int. Electron Devices Meet. IEDM*, p. 9.5.1-9.5.4, 2011.

[13] W.Y. Chang *et al.*, "Improvement of resistive switching characteristics in TiO2 thin films with embedded Pt nanocrystals," *Appl. Phys. Lett.*, vol. 95, no. 4, p. 42104, 2009.

[14] D.-T. Wang *et al.*, "Resistive Switching and Synaptic Behaviors of TaN/Al2O3/ZnO/ITO Flexible Devices With Embedded Ag Nanoparticles," *IEEE Electron Device Lett.*, vol. 37, no. 7, pp. 878–881, 2016.

[15] C. A. Pan and T. P. Ma, "Work function of In2O3 film as determined from internal photoemission," *Appl. Phys. Lett.*, vol. 37, no. 8, pp. 714–716, 1980.